

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平 1 0 - 2 2 2 4 1 1

(43) 公開日 平成 1 0 年 (1 9 9 8) 8 月 2 1 日

(51) Int. Cl.
G06F 12/00

識別記号 庁内整理番号
545

F I
G06F 12/00

技術表示箇所
545 Z

審査請求 未請求 請求項の数 3 O L (全 6 頁)

(21) 出願番号 特願平 9 - 2 3 0 1 3

(22) 出願日 平成 9 年 (1 9 9 7) 2 月 5 日

(71) 出願人 0 0 0 0 0 4 2 2 6

日本電信電話株式会社

東京都新宿区西新宿三丁目 1 9 番 2 号

(72) 発明者 朝香 卓也

東京都新宿区西新宿三丁目 1 9 番 2 号 日

本電信電話株式会社内

(74) 代理人 弁理士 伊東 忠彦

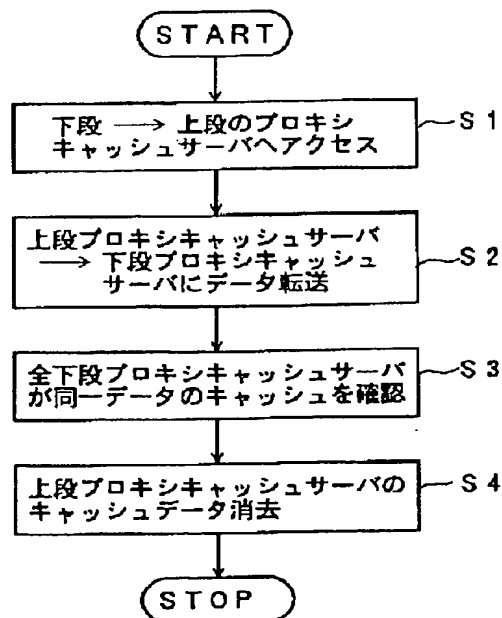
(54) 【発明の名称】 プロキシキャッシュサーバ制御方法及びプロキシキャッシュサーバ

(57) 【要約】

【課題】 上段プロキシキャッシュサーバには問い合わせがある可能性のあるデータだけが残し、キャッシュ可能データ容量を有効に利用することができ、結果として、キャッシュデータのヒット率を上げることが可能なプロキシキャッシュサーバ制御方法及びプロキシキャッシュサーバを提供することを目的とする。

【解決手段】 本発明は、多段プロキシキャッシュサーバのうち、上段プロキシキャッシュサーバにおいて、上段プロキシキャッシュサーバの直下に位置する全ての下段プロキシキャッシュサーバの全てが、同一のデータをキャッシュしたことが確認された時点で、上段のプロキシキャッシュサーバのデータを削除する。

本発明の原理を説明するための図



BEST AVAILABLE COPY

【特許請求の範囲】

【請求項 1】 WWW等の蓄積型データ転送を目的とするネットワークの多段プロキシキャッシュサーバネットワークにおいて、プロキシキャッシュサーバを制御するプロキシキャッシュサーバ制御方法において、

多段プロキシキャッシュサーバのうち、上段プロキシキャッシュサーバの直下に位置する全ての下段プロキシキャッシュサーバが、ある同一のデータをキャッシュしたことが確認された時点で、上段のプロキシキャッシュサーバの該データを削除することを特徴とするプロキシキャッシュサーバ制御方法。

【請求項 2】 前記多段プロキシキャッシュサーバのうち、ユーザホストに接続される下段プロキシキャッシュサーバから前記上段プロキシキャッシュサーバにアクセスし、

前記上段プロキシキャッシュサーバから前記下段プロキシキャッシュサーバに対してキャッシュデータを転送し、

前記上段プロキシキャッシュサーバは、すべての下段プロキシキャッシュサーバが同一のデータをキャッシュしたことを確認した時点で、該上段プロキシキャッシュサーバのキャッシュデータを消去する請求項 1 記載のプロキシキャッシュサーバ制御方法。

【請求項 3】 WWW等の蓄積型データ転送を目的とするネットワーク上の多段プロキシキャッシュサーバから構成され、該プロキシキャッシュサーバを制御するプロキシキャッシュサーバであって、

自プロキシキャッシュサーバより下段のプロキシキャッシュサーバからのアクセスに基づいてキャッシュデータを転送するデータ転送手段と、

自プロキシキャッシュサーバより下段のプロキシキャッシュサーバの全てが、前記キャッシュデータと同一のキャッシュデータをキャッシュしたか否かを判定するキャッシュ判定手段と、

前記キャッシュ判定手段において、キャッシュされたことが判定されると、上段のプロキシキャッシュサーバのデータを削除するキャッシュデータ削除手段とを有することを特徴とするプロキシキャッシュサーバ。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】 本発明は、プロキシキャッシュサーバ制御方法及びプロキシキャッシュサーバに係り、特に、WWW等の蓄積型データ転送を目的とするネットワークにおいて、プロキシキャッシュサーバに蓄積されるキャッシュデータを効率よく蓄積するための多段に接続されたプロキシキャッシュサーバの制御方法及びプロキシキャッシュサーバに関する。

【 0 0 0 2 】

【従来の技術】 各ユーザは、プロキシキャッシュサーバを経由し、目的とするデータを持つWWWページ等のデ

ータを配備しているデータ蓄積サーバにアクセスする。あるユーザがあるデータ蓄積サーバにアクセスし、データ蓄積サーバからユーザ側へデータ転送を行うときに、プロキシキャッシュサーバは、当該データをキャッシュする。他のユーザが当該データを要求するときに、プロキシキャッシュサーバは、データ蓄積サーバのデータの代わりに、キャッシュされたデータをユーザに転送する。このプロキシキャッシュサーバが図 6 に示すように、多段接続されているなら、上段のプロキシキャッシュサーバは、複数の下段プロキシキャッシュサーバがキャッシュしているデータ全てをキャッシュすることができる。あるユーザがあるデータを取得しようとした時、下段プロキシキャッシュサーバがキャッシュしていなかったとしても、上段プロキシキャッシュサーバがデータをキャッシュしていれば、当該ユーザは、上段プロキシキャッシュサーバからデータを取得することができる。結果として、キャッシュデータのヒット率を上げることができる。

【 0 0 0 3 】 従来のツリー上に多段接続されたプロキシキャッシュサーバでは、各プロキシキャッシュサーバは、独立にデータのキャッシュ及びキャッシュ消去を行っている図 6 の構成における上段及び下段のプロキシキャッシュサーバがキャッシュしていなかったら、あるユーザがデータ蓄積サーバのデータ取得時において、上段及び下段のプロキシキャッシュサーバは、共に当該データをキャッシュする。キャッシュデータを消去する場合には、各プロキシキャッシュサーバは、「現在時刻－データのキャッシュ時刻」と閾値を比較して、閾値を越えていれば、消去する。あるいは、新しいデータをキャッシュするときに、プロキシキャッシュサーバのキャッシュ可能データ容量を越えるなら、キャッシュデータ置き換えアルゴリズムに従い消去される。置き換えアルゴリズムとしては、例えば、キャッシュ時刻の古いものあるいは、データ量の大きなものを消去するといったものがある。

【 0 0 0 4 】 以上をツリー上に多段接続されたプロキシキャッシュサーバの処理をまとめると以下ようになる。なお、従来のプロキシキャッシュサーバでは、上下段のプロキシキャッシュサーバともに同じ処理を行う。

〈データ蓄積時〉

・下段・上段プロキシキャッシュサーバ共に当該データをキャッシュしていなかった場合：上段及び下段のプロキシキャッシュサーバとともにキャッシュすると同時に、当該データをユーザに転送する。

【 0 0 0 5 】 ・上段プロキシキャッシュサーバだけが当該データキャッシュしていた場合：下段プロキシキャッシュサーバがキャッシュすると同時に、当該データをユーザに転送する。

・下段が当該データをキャッシュしていた場合：下段プロキシキャッシュサーバが、当該データをユーザに転送

10

20

30

40

50

する。

【 0 0 0 6 】 〈データ消去時〉

・データのキャッシュされてからの経過時間>閾値となるキャッシュデータがある場合：当該データを消去する。

・キャッシュ可能データ容量を越えて新しくデータをキャッシュしようとする場合：キャッシュデータ置き換えアルゴリズムに従い消去する。

【 0 0 0 7 】

【発明が解決しようとする課題】しかしながら、上記従来のツリー上に多段接続されたプロキシキャッシュサーバでは、上段のプロキシキャッシュサーバのデータ消去時、特に、プロキシキャッシュサーバのキャッシュ可能データ容量を越える場合に、キャッシュ時刻の古いもの、あるいは、データ量の大きなものを消去することは効率的でない。なぜなら、多段接続されたプロキシキャッシュサーバでは、下段プロキシキャッシュサーバ全てが、キャッシュしているデータは、もはや、上段プロキシキャッシュサーバに対する問い合わせがないにもかかわらず、上段プロキシキャッシュサーバが当該データをキャッシュデータとして保持し続けることによる。

【 0 0 0 8 】本発明は、上記の点に鑑みなされたもので、従来の上段プロキシキャッシュサーバにおける無駄なキャッシュデータ保持をなくすことによって、上段プロキシキャッシュサーバには問い合わせがある可能性のあるデータだけが残し、キャッシュ可能データ容量を有効に利用することができ、結果として、キャッシュデータのヒット率を上げることが可能なプロキシキャッシュサーバ制御方法及びプロキシキャッシュサーバを提供することを目的とする。

【 0 0 0 9 】

【課題を解決するための手段】図 1 は、本発明の原理を説明するための図である。本発明は、WWW等の蓄積型データ転送を目的とするネットワークの多段プロキシキャッシュサーバネットワークにおいて、プロキシキャッシュサーバを制御するプロキシキャッシュサーバ制御方法において、多段プロキシキャッシュサーバのうち、上段プロキシキャッシュサーバの直下に位置する全ての下段プロキシキャッシュサーバが、ある同一のデータをキャッシュしたことが確認された時点で、上段のプロキシキャッシュサーバの該データを削除する。

【 0 0 1 0 】また、本発明は、多段プロキシキャッシュサーバのうち、ユーザホストに接続される下段プロキシキャッシュサーバから上段プロキシキャッシュサーバにアクセスし（ステップ 1）、上段プロキシキャッシュサーバから下段プロキシキャッシュサーバに対してキャッシュデータを転送し（ステップ 2）、上段プロキシキャッシュサーバは、すべての下段プロキシキャッシュサーバが同一のデータをキャッシュしたことを確認した時点で（ステップ 3）、該上段プロキシキャッシュサーバの

キャッシュデータを消去する（ステップ 4）。

【 0 0 1 1 】図 2 は、本発明の原理構成図である。本発明は、WWW等の蓄積型データ転送を目的とするネットワーク上の多段プロキシキャッシュサーバ 1 0 0、2 0 0 から構成され、該プロキシキャッシュサーバ 1 0 0、2 0 0 を制御するプロキシキャッシュサーバであって、自プロキシキャッシュサーバ 1 0 0 より下段のプロキシキャッシュサーバ 2 0 0 からのアクセスに基づいてキャッシュデータを転送するデータ転送手段 1 0 と、自プロキシキャッシュサーバ 1 0 0 より下段の全てのプロキシキャッシュサーバ 2 0 0 が、同一のキャッシュデータをキャッシュしたか否かを判定するキャッシュ判定手段 2 0 と、キャッシュ判定手段 2 0 において、キャッシュされたことが判定されると、自プロキシキャッシュサーバ 1 0 0 のデータを削除するキャッシュデータ削除手段 3 0 とを有する。

【 0 0 1 2 】上記のように、本発明では、アクセスのあった下段プロキシキャッシュサーバにキャッシュデータを転送し、当該キャッシュデータがキャッシュされたことを確認し、その確認がとれた時点で、転送元のプロキシキャッシュサーバのキャッシュデータを削除することにより、上段プロキシキャッシュサーバに、アクセスされる可能性のあるデータのみが残されることになる。

【 0 0 1 3 】

【発明の実施の形態】図 3 は、本発明のプロキシキャッシュサーバの構成を示す。同図に示すネットワークの図は、WWWページ等の蓄積型データ転送を目的とするネットワークを表す。同図に示す多段プロキシキャッシュサーバは、上段プロキシキャッシュサーバ 1、下段プロキシキャッシュサーバ 2、3 及び、WWWページの取得を要求するユーザのホスト 6 から構成される。

【 0 0 1 4 】上段プロキシキャッシュサーバ 1 は、キャッシュデータ管理部 4 と下段プロキシキャッシュサーバ情報管理部 5 を有する。キャッシュデータ管理部 4 は、WWWサーバのWWWページのデータを蓄積し、管理する。下段プロキシキャッシュサーバ情報管理部 5 は、どのキャッシュデータをどの下段プロキシキャッシュサーバが保持しているかという情報を管理する。

【 0 0 1 5 】上段プロキシキャッシュサーバ 1 に対して下段プロキシキャッシュサーバ A 2 からキャッシュデータの要求があると、上段プロキシキャッシュサーバ 1 のキャッシュデータ管理部 4 は、蓄積しているキャッシュデータを下段プロキシキャッシュサーバ A 2 に転送すると共に、キャッシュデータを下段プロキシキャッシュサーバ A 2 に転送したことを下段プロキシキャッシュサーバ情報管理部 5 に通知する。これにより、下段プロキシキャッシュサーバ情報管理部 5 は、キャッシュデータの転送先を管理する。

【 0 0 1 6 】なお、上段プロキシキャッシュサーバ 1、下段プロキシキャッシュサーバ A 2、下段プロキシキャッシュサーバ

10

20

30

40

50

ッシュサーバ B 3 共に、従来のキャッシュ制御方法に加えて、本発明による上段プロキシキャッシュサーバ 1 のキャッシュデータの消去も行。図 4 は、本発明の上段プロキシキャッシュサーバのキャッシュデータ消去動作のシーケンスチャートである。

【0017】ステップ 101) ユーザホスト 6 は、下段プロキシキャッシュサーバ 2 へアクセスする。

ステップ 102) 下段プロキシキャッシュサーバ 2 において望む WWW ページのデータがなく、下段プロキシキャッシュサーバ 2 は、上段プロキシキャッシュサーバ 1 へアクセスする。

【0018】ステップ 103) 上段プロキシキャッシュサーバ 1 は、キャッシュデータを下段プロキシキャッシュサーバ 2 に対して転送する。

ステップ 104) 既に、下段プロキシキャッシュサーバ 2、3 (図 3 の例では、下段プロキシキャッシュサーバ A、B) が同一のデータをキャッシュすることを上段プロキシキャッシュサーバ 1 内の下段プロキシキャッシュサーバ情報管理部 5 が検出したら、上段プロキシキャッシュサーバ 1 内のキャッシュデータ管理部 4 にある当該キャッシュデータを消去する。

【0019】

【実施例】以下、本発明の実施例を前述の図 3 及び図 4 に基づいて説明する。上記の一連の動作を図 3 の例に適用すると、上段プロキシキャッシュサーバ 1 には、当該プロキシキャッシュサーバ 1 の下段に位置する下段プロキシキャッシュサーバ A 2 と下段プロキシキャッシュサーバ B 3 のみが接続されている。

【0020】ここで、ユーザホスト 6 からデータ要求された下段プロキシキャッシュサーバ A 2 から上段プロキシキャッシュサーバ 1 に対してデータ要求を発行すると、上段プロキシキャッシュサーバ 1 のキャッシュデータ管理部 4 は、下段プロキシキャッシュサーバ A 2 に対してキャッシュデータを転送すると共に、転送した下段プロキシキャッシュサーバ A 2 の情報 (=A) 及びキャッシュデータ識別子 (ID=a) を、下段プロキシキャッシュサーバ情報管理部 5 に通知する。これにより、下段プロキシキャッシュサーバ情報管理部 5 は、転送したサーバ名「A」とキャッシュデータ識別子 (ID=a)、及び時刻等を記録する。

【0021】次に、下段プロキシキャッシュサーバ B 3 から上段プロキシキャッシュサーバ 1 に下段プロキシキャッシュサーバ A 2 から要求されたデータと同様のデータ要求が発行されると、上段プロキシキャッシュサーバ 1 のキャッシュデータ管理部 4 は、当該データ要求に対応するキャッシュデータを下段プロキシキャッシュサーバ B 3 に転送する。そして、下段プロキシキャッシュサーバ情報管理部 5 に対して、データ転送した下段プロキシキャッシュサーバ B 3 の情報 (=B) 及びキャッシュデータ識別子 (ID=a) を下段プロキシキャッシュサ

サーバ情報管理部 5 に通知する。これにより、下段プロキシキャッシュサーバ情報管理部 5 は、転送したサーバ名「B」とキャッシュデータ識別子 (ID=a)、及び時刻等を記録する。

【0022】上段プロキシキャッシュサーバ 1 に接続されている下段プロキシキャッシュサーバは、下段プロキシキャッシュサーバ A 2 と下段プロキシキャッシュサーバ B 3 のみであるため、下段プロキシキャッシュサーバ情報管理部 5 は、すべての下段プロキシキャッシュサーバ A 2、下段プロキシキャッシュサーバ B 3 に対して同一のキャッシュデータ (ID=a) が転送されたことを確認する。これにより、キャッシュデータ管理部 4 に格納されている上記キャッシュデータ識別子 (ID=a) のキャッシュデータを消去する。

【0023】なお、本発明は、上記の実施例に限定されことなく、特許請求の範囲内で種々変更・応用が可能である。

【0024】

【発明の効果】以下に、本発明の効果を説明する。まず、図 6 のモデルを用いて説明する。同図において、上段プロキシキャッシュサーバを 1 台、下段プロキシキャッシュサーバを 2 台、総ユーザ数を 200 のもとで、各下段プロキシキャッシュサーバにそれぞれ 100 ユーザずつがアクセスするものとする。各ユーザからのデータへのアクセスは、ポアソン到着するものとし、その平均アクセス時間間隔を 100 とした。ここでは、データ i へのアクセスは、

$$\lambda(i) = q(1 - q)^i \quad (1)$$

に従い発生するものとした。但し、q は、アクセス先データのインデックス i の平均を決定するパラメータで、その平均値は $1/q$ によって得られる。ここでは、 $q = 0.99$ とした。これにより、i が小さいほどアクセス率が高いことになる。また、図 6 でのユーザは、データ蓄積サーバは持っていないものとする。500 単位時間でタイムアウトし、タイムアウトしたキャッシュデータは消去される。また、プロキシキャッシュサーバ内のキャッシュデータ数は、プロキシキャッシュサーバのエントリ数を上限とし、データ量による上限は考えない。また、エントリ数分のキャッシュデータが既にある状態で、新しいデータがキャッシュされる場合には、従来方法では、置き換えアルゴリズム LRU (Least Recently Used) に従い、古いデータが消去される。置き換えアルゴリズム LRU とは、できるだけ最終アクセス時刻が古いキャッシュデータから消去される。また、上段・下段ともプロキシキャッシュサーバのエントリ数上限は同一とした。

【0025】図 5 は、本発明の効果を示すための評価結果である。各プロキシキャッシュサーバがキャッシュできるエントリ数が増えた場合のエントリあたりのヒット率を求めた。同図の評価結果からわかるように、上段

プロキシキャッシュサーバのヒット率について、提案方法は、従来方法に比べて有効であることが判る。

【0026】本発明の方法による上段プロキシキャッシュサーバのキャッシュデータ消去手順は、従来からのキャッシュ制御方法と併用することが可能である。その理由は、本発明の方法における消去実行時点が従来の消去実行時点と異なることによる。また、本発明では、上段プロキシキャッシュサーバだけの機能として実現できる。上段プロキシキャッシュサーバは下段プロキシキャッシュサーバのキャッシュするデータを知ることが常に
10 できることから、本発明では、上段プロキシキャッシュサーバと下段プロキシキャッシュサーバ間で特別なプロトコルは必要としない。

【図面の簡単な説明】

【図1】本発明の原理を説明するための図である。

【図2】本発明の原理構成図である。

【図3】本発明のプロキシキャッシュサーバの構成図である。

【図4】本発明の上段プロキシキャッシュサーバのキャッシュデータ消去動作のシーケンスチャートである。

【図5】本発明の効果を示すための評価結果である。

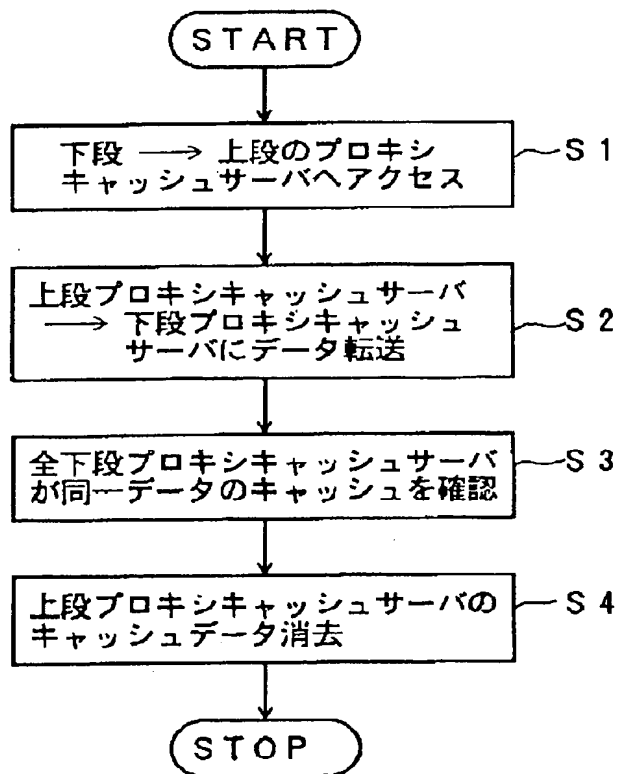
【図6】多段接続されているプロキシキャッシュサーバの例である。

【符号の説明】

- 1 上段プロキシキャッシュサーバ
- 2 下段プロキシキャッシュサーバA
- 3 下段プロキシキャッシュサーバB
- 4 キャッシュデータ管理部
- 5 下段プロキシキャッシュサーバ情報管理部
- 6 ユーザホスト
- 10 データ転送手段
- 20 キャッシュ判定手段
- 30 キャッシュデータ削除手段
- 100 上段プロキシキャッシュサーバ
- 200 下段プロキシキャッシュサーバ

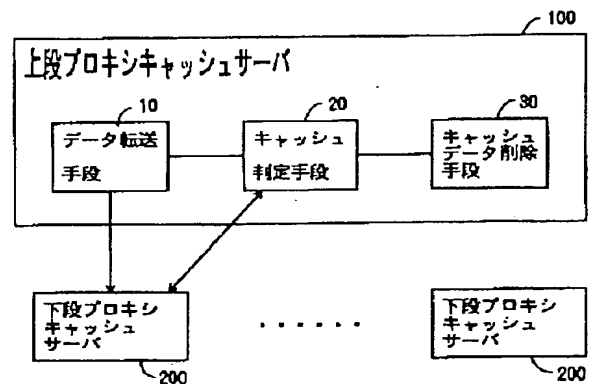
【図1】

本発明の原理を説明するための図



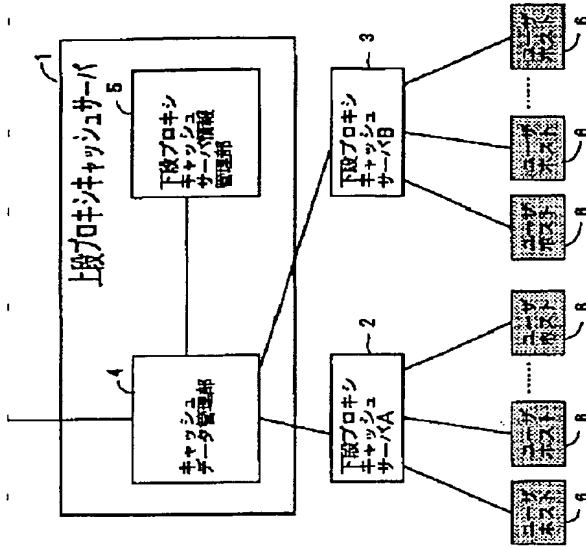
【図2】

本発明の原理構成図



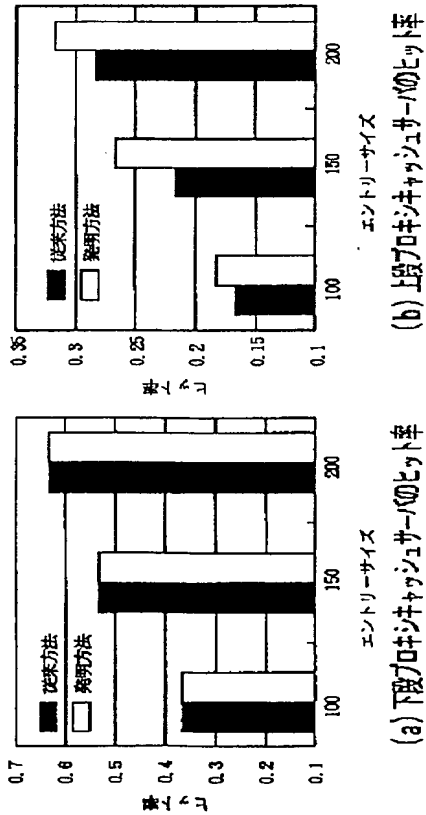
【 図 3 】

本発明のプロキシキャッシュサーバの構成図

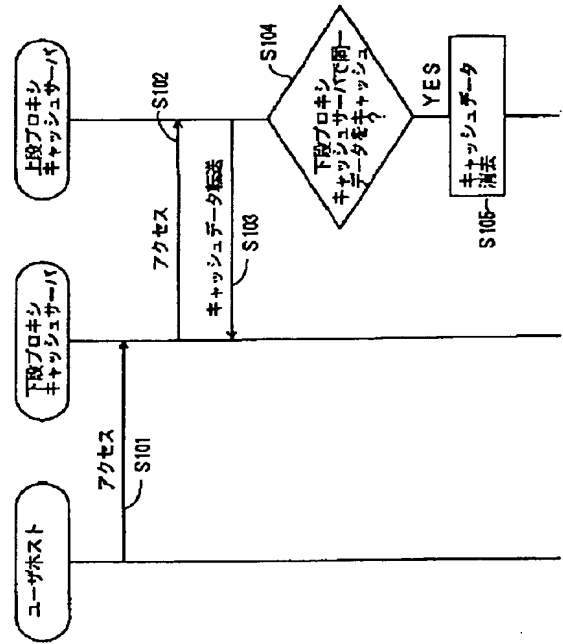


【 図 5 】

本発明の効果をj示すためのj評価結果

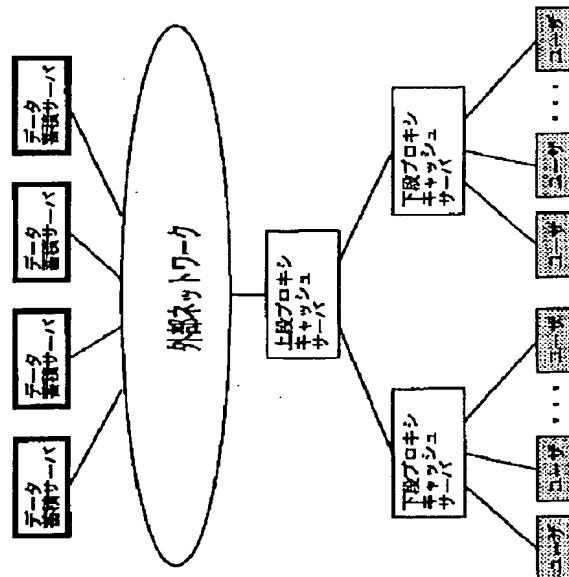


【 図 4 】

本発明の上段プロキシキャッシュサーバの
キャッシュデータ消去動作のシーケンスチャート

【 図 6 】

多段j接続されているプロキシキャッシュサーバの例



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.